



Master Thesis

Specialize *without* Forgetting: Retaining Generalist Skills *in* Fine-Tuned VLAs

Univ.-Prof. Dr.-Ing. Volkmar Schulz
Head of Institute

Dr.-Ing. Yuli Wu
Group Leader

Physical AI Lab
Chair of Imaging and Computer Vision
RWTH Aachen University
Kopernikusstr. 16
52074 Aachen, Germany

<https://pai.rwth-aachen.de/>

*We are inviting applications for the **Master Thesis Topic** in the **Physical AI Lab** at the Chair of Imaging and Computer Vision, Faculty of Electrical Engineering and Information Technology, RWTH Aachen University. Our group develops generalist robots that autonomously solve real-world long-horizon tasks.*

Introduction of the Topic

Vision-language-action (VLA) models such as $\pi 0.5$ [1] combine internet-scale vision-language pretraining with large robot datasets to perform diverse manipulation tasks from language instructions. To reach deployment-level reliability, practitioners typically fine-tune them on in-house demonstrations for a specific task. This yields strong target-task performance but causes catastrophic forgetting [2] of previously acquired skills. Since open-weight VLAs are often released without their pretraining data, rehearsing on that data is not an option. RETAIN [3] offers a data-free alternative by interpolating the fine-tuned and pretrained weights, recovering generalist performance without extra training. However, its coarse, module-level coefficients must be tuned through costly robot evaluations. Based on RETAIN [3], this thesis investigates if finer-grained, data-free strategies can better balance learning a new task with retaining pretrained skills.

[1] K. Black et al., " $\pi 0.5$: A Vision-Language-Action Model with Open-World Generalization," CoRL, 2025.

[2] J. Kirkpatrick et al., "Overcoming Catastrophic Forgetting in Neural Networks," PNAS, 2017.

[3] Y. Yadav et al., "Robust Fine-Tuning of Vision-Language-Action Robot Policies via Parameter Merging," ICLR, 2026.

Your Profile

- Background in machine learning, robot learning, computer vision, or related fields
- Strong Python and PyTorch skills
- High motivation and the ability to work independently and in a team
- Interest in embodied AI and vision-language-action models
- A plus: experience with model fine-tuning, robot simulators (e.g., LIBERO, MuJoCo), or real robots

What We Offer

- An open research question rather than a reimplement task
- Close supervision and strong support for publications
- A clear starting point: an open-weight VLA ($\pi 0.5$), the RETAIN baseline, and established benchmarks
- Access to on-premises GPUs, high-end GPU clusters, and real robots

Application

Duration of the master thesis is 6 months. Please send your application to the following email address yuli.wu@lfb.rwth-aachen.de using the subject "**NoForget_{Name}**", and include your CV and academic transcripts.